import torch

import pandas as pd

from transformers import GPT2LMHeadModel, GPT2Tokenizer, TrainingArguments, Trainer

# Initialize the GPT-2 tokenizer and model

tokenizer = GPT2Tokenizer.from\_pretrained("gpt2")

model = GPT2LMHeadModel.from\_pretrained("gpt2")

def generate\_flavor\_profile(molecule\_a, molecule\_b):

# Combine the two molecules and create a prompt

prompt = f"Molecule A is related to {molecule\_a} and Molecule B is related to {molecule\_b}. Taste Description is:"

# Tokenize the prompt

input\_ids = tokenizer.encode(prompt, return\_tensors="pt")

# Generate text using the GPT-2 model with attention\_mask and pad\_token\_id set

with torch.no\_grad():

output = model.generate(input\_ids, max\_length=50, num\_return\_sequences=1,

pad\_token\_id=tokenizer.eos\_token\_id,

no\_repeat\_ngram\_size=2,

top\_k=40,

top\_p=0.90,

temperature=0.8,

do\_sample=True)

# Decode the generated text

generated\_text = tokenizer.decode(output[0], skip\_special\_tokens=True)

# Extract just the taste description part

taste\_description = generated\_text[len(prompt):].strip()

return taste\_description

# Example usage

molecule\_a = "Vanillic Acid"

molecule\_b = "Vanillin"

generated\_taste = generate\_flavor\_profile(molecule\_a, molecule\_b)

print(f"Taste Description for {molecule\_a} and {molecule\_b}: {generated\_taste}")